REMARKS

Claims 21-45 are currently pending in this application of which claims 23-31 are actively being prosecuted. Claims 21, 22, and 32-45 are withdrawn from consideration as being drawn to non-elected inventions.

Claims 21, 30, and 31 have been amended. To expedite prosecution, claim 21 has been amended to recite that the claimed variants and biologically active fragments have alcohol dehydrogenase activity, and claim 30 has been amended to recite that the claimed polynucleotide variants encode a polypeptide having alcohol dehydrogenase activity. Support for the amendment of claims 21 and 30 can be found in the specification, for example, at pages 1-2, which describe alcohol dehydrogenases, page 14, lines 3-17, which points out the presence of alcohol dehydrogenase motifs in ScrM-1, at page 46, which describe assays for dehydrogenase activity, and Example XII, which describes immunogenic fragments. In addition, a recent Blast analysis (Exhibit A) shows that SEQ ID NO:1 is 96% identical to the human peroxisomal short-chain alcohol dehydrogenase (g12804321). By this amendment, Applicants expressly do not disclaim equivalents of the invention which could include polypeptides having biological activities in addition to the recited alcohol dehydrogenase activity or polynucleotides encoding such polypeptides. These amendments further clarify the intended subject matter of the claimed invention and address the rejections under 35 U.S.C. § 112, first and second paragraphs, and 35 U.S.C. § 102. Claim 31 has been amended to recite an isolated polynucleotide consisting of at least 60 contiguous nucleotides of a polynucleotide of SEQ ID NO:3 in order to address the written description rejection under 35 U.S.C. § 112, first paragraph. Entry of these amendments is respectfully requested.

Applicants reserve the right to prosecute non-elected subject matter in subsequent divisional applications.

Comments Regarding Restriction Requirement

Applicants affirm the election with traverse of Group III, which corresponds to claims 23-31, drawn to polynucleotides.

Applicants reiterate the request that the Examiner withdraw the Restriction Requirement at least with respect to claims 21, 22, 35, and 36 of Group I, and examine those claims together with the elected polynucleotide claims of Group III.

The rules under MPEP section 1893.03(d) require the Examiner to apply the Unity of Invention standard PCT Rule 13.2 instead of U.S. restriction/election of species practice in national stage applications, such as the instant application filed under 35 U.S.C. 371. Applicants submit that unity of invention exists for claims drawn to the polypeptide sequence of SEQ ID NO:1 (*i.e.,* claims 21, 22, 35, and 36) and claims drawn to the elected polynucleotide sequence of SEQ ID NO:3 which encodes SEQ ID NO:1 (*i.e.,* claims 23-25) based on the rules concerning unity of invention under the Patent Cooperation Treaty. The Administrative Instructions Under The Patent Cooperation Treaty, Annex B, Unity of Invention, Part 2, "Examples Concerning Unity of Invention" provide the following guidelines with regard to unity of invention between a protein and the polynucleotide that encodes it:

> *Example 17*
> Claim 1: Protein X.
> Claim 2: DNA sequence encoding protein X.
> Expression of the DNA sequence in a host results in the production of a protein which is determined by the DNA sequence. The protein and the DNA sequence exhibit corresponding special technical features. Unity between claims 1 and 2 is accepted.

Applicants submit that Example 17 does apply to the claims of the instant application, since the polynucleotide of SEQ ID NO:3 does encode the polypeptide of SEQ ID NO:1. In particular, claims 21 and 23 meet the unity of invention standard. Claim 23 recites an isolated polynucleotide encoding a polypeptide of claim 21. Unity of invention is accepted between a protein and the polynucleotide that encodes it. The refusal to examine claims drawn to polynucleotides and polypeptides together on the grounds that the polynucleotide is structurally distinct from the polypeptide is improper.

Rejoinder

Applicants request that claims 32-34 and 39-42, drawn to methods of using the elected polynucleotides of Group III, be rejoined per the Commissioner's Notice in the Official Gazette of March 26, 1996, entitled "Guidance on Treatment of Product and Process Claims in light of *In re Ochiai, In re Brouwer* and 35 U.S.C. § 103(b), which sets forth the rules that upon allowance

of any of the product claims, the method claims covering the same scope of products be rejoined. Applicants request that claims 32-34 and 39-42 be rejoined and examined upon allowance of any claim drawn to the claimed polynucleotides.

Objections to the claims

Claims 23-30 are objected to because of their dependence from a non-elected base claim. As mentioned above, Applicants believe that the claims drawn to the polypeptides of the invention, according to the unity of invention standard, should be examined with the elected claims drawn to the polynucleotides currently under examination. Applicants request reconsideration and believe amending the claims at this time would be premature.

Written description rejections under 35 U.S.C. § 112, first paragraph

Claims 23 and 26-31 have been rejected under the first paragraph of 35 U.S.C. 112 for alleged lack of an adequate written description. This rejection is respectfully traversed.

The requirements necessary to fulfill the written description requirement of 35 U.S.C. 112, first paragraph, are well established by case law.

> . . . the applicant must also convey with reasonable clarity to those skilled in the art that, as of the filing date sought, he or she was in possession *of the invention*. The invention is, for purposes of the "written description" inquiry, *whatever is now claimed. Vas-Cath, Inc. v. Mahurkar*, 19 USPQ2d 1111, 1117 (Fed. Cir. 1991)

Attention is also drawn to the Patent and Trademark Office's own "Guidelines for Examination of Patent Applications Under the 35 U.S.C. Sec. 112, para. 1", published January 5, 2001, which provide that :

> An applicant may also show that an invention is complete by disclosure of sufficiently detailed, relevant identifying characteristics which provide evidence that applicant was in possession of the claimed invention, i.e., complete or partial structure, other physical and/or chemical properties, functional characteristics when coupled with a known or disclosed correlation between function and structure, or some combination of such characteristics. What is conventional or well known to one of ordinary skill in the art need not be disclosed in detail. If a skilled artisan would have understood the inventor to be in possession of the claimed invention at the time of filing, even if every nuance of the claims is not explicitly described in the specification, then the adequate description requirement is met. (footnotes omitted.)

Thus, the written description standard is fulfilled by both what is specifically disclosed and what is conventional or well known to one skilled in the art.

SEQ ID NO:1 and SEQ ID NO:3 are specifically disclosed in the application (see, for example, page 2, lines 31-34, page 3, lines 18-20). Variants of SEQ ID NO:3 are described, for example, at page 15, lines 19-29. Incyte clones in which the nucleic acids encoding the human ScRM-1 were first identified and libraries from which those clones were isolated are described, for example, at page 13, lines 28-33 of the Specification. Chemical and structural features of SEQ ID NO:1 are described, for example, on page 13, line 34 through page 14, line 22. Given SEQ ID NO:1 and SEQ ID NO:3, one of ordinary skill in the art would recognize a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence of SEQ ID NO:3, a polynucleotide encoding a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence of SEQ ID NO:1, a polynucleotide encoding a fragment of SEQ ID NO:1, or a fragment of the polynucleotide of SEQ ID NO:3. Accordingly, the Specification provides an adequate written description of the recited polynucleotide sequences.

The Office Action has further asserted that the claims are not supported by an adequate written description because "the claims are drawn to a large variable genus of polynucleotides encoding polypeptides having unknown activity or inactive variants" (Office Action, page 4).

Such a position is believed to present a misapplication of the law.

1. The present claims specifically define the claimed genus through the recitation of chemical structure

Court cases in which "DNA claims" have been at issue commonly emphasize that the recitation of structural features or chemical or physical properties are important factors to consider in a written description analysis of such claims. For example, in *Fiers v. Revel*, 25 USPQ2d 1601, 1606 (Fed. Cir. 1993), the court stated that:

> If a conception of a DNA requires a precise definition, such as by structure, formula, chemical name or physical properties, as we have held, then a description also requires that degree of specificity.

In a number of instances in which claims to DNA have been found invalid, the courts

have noted that the claims attempted to define the claimed DNA in terms of functional characteristics without any reference to structural features. As set forth by the court in *University of California v. Eli Lilly and Co.*, 43 USPQ2d 1398, 1406 (Fed. Cir. 1997):

> In claims to genetic material, however, a generic statement such as "vertebrate insulin cDNA" or "mammalian insulin cDNA," without more, is not an adequate written description of the genus because it does not distinguish the claimed genus from others, except by function.

Thus, the mere recitation of functional characteristics of a DNA, without the definition of structural features, has been a common basis by which courts have found invalid claims to DNA. For example, in *Lilly*, 43 USPQ2d at 1407, the court found invalid for violation of the written description requirement the following claim of U.S. Patent No. 4,652,525:

> 1. A recombinant plasmid replicable in procaryotic host containing within its nucleotide sequence a subsequence having the structure of the reverse transcript of an mRNA of a vertebrate, which mRNA encodes insulin.

In *Fiers*, 25 USPQ2d at 1603, the parties were in an interference involving the following count:

> A DNA which consists essentially of a DNA which codes for a human fibroblast interferon-beta polypeptide.

Party Revel in the *Fiers* case argued that its foreign priority application contained an adequate written description of the DNA of the count because that application mentioned a potential method for isolating the DNA. The Revel priority application, however, did not have a description of any particular DNA structure corresponding to the DNA of the count. The court therefore found that the Revel priority application lacked an adequate written description of the subject matter of the count.

Thus, in *Lilly* and *Fiers*, nucleic acids were defined on the basis of functional characteristics and were found not to comply with the written description requirement of 35 U.S.C. §112; *i.e.*, "an mRNA of a vertebrate, which mRNA encodes insulin" in *Lilly*, and "DNA which codes for a human fibroblast interferon-beta polypeptide" in *Fiers*. In contrast to the situation in *Lilly* and *Fiers*, the claims at issue in the present application define polynucleotides in terms of chemical structure, rather than functional characteristics. For example, the "variant language" of independent claim 30 recites chemical structure to define the claimed genus:

An isolated polynucleotide selected from the group consisting of:...

b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence of SEQ ID NO:3, said polynucleotide encoding a polypeptide having alcohol dehydrogenase activity...

From the above it should be apparent that the claims of the subject application are fundamentally different from those found invalid in *Lilly* and *Fiers*. The subject matter of the present claims is defined in terms of the chemical structures of SEQ ID NO:1 and SEQ ID NO:3. In the present case, there is no reliance merely on a description of functional characteristics of the polynucleotides recited by the claims. Moreover, functional recitations add to the structural characterization of the recited polynucleotides. The polynucleotides defined in the claims of the present application recite structural features, and cases such as *Lilly* and *Fiers* stress that the recitation of structure is an important factor to consider in a written description analysis of claims of this type. By failing to base its written description inquiry "on whatever is now claimed," the Office Action failed to provide an appropriate analysis of the present claims and how they differ from those found not to satisfy the written description requirement in *Lilly* and *Fiers*

2. The present claims do not define a genus which is "highly variant"

Furthermore, the claims at issue do not describe a genus which could be characterized as "highly variant." Available evidence illustrates that the claimed genus is of narrow scope.

In support of this assertion, the Examiner's attention is directed to the enclosed reference by Brenner et al. ("Assessing sequence comparison methods with reliable structurally identified distant evolutionary relationships," Proc. Natl. Acad. Sci. USA (1998) 95:6073-6078). Through exhaustive analysis of a data set of proteins with known structural and functional relationships and with <90% overall sequence identity, Brenner et al. have determined that 30% identity is a reliable threshold for establishing evolutionary homology between two sequences aligned over at least 150 residues. (Brenner et al., pages 6073 and 6076.) Furthermore, local identity is particularly important in this case for assessing the significance of the alignments, as Brenner et al. further report that ≥40% identity over at least 70 residues is reliable in signifying homology between proteins. (Brenner et al., page 6076.)

The present application is directed, *inter alia*, to alcohol dehydrogenases related to the amino acid sequence of SEQ ID NO:1. In accordance with Brenner et al, naturally occurring molecules may exist which could be characterized as alcohol dehydrogenases and which have as little as 40% identity over at least 70 residues to SEQ ID NO:1. The "variant language" of the present claims recites, for example, polynucleotides encoding "a naturally-occurring amino acid sequence having at least 90% sequence identity to the sequence of SEQ ID NO:1" (note that SEQ ID NO:1 has 278 amino acid residues). This variation is far less than that of all potential alcohol dehydrogenases related to SEQ ID NO:1, i.e., those alcohol dehydrogenases having as little as 40% identity over at least 70 residues to SEQ ID NO:1.

3. The state of the art at the time of the present invention is further advanced than at the time of the *Lilly* and *Fiers* applications

In the *Lilly* case, claims of U.S. Patent No. 4,652,525 were found invalid for failing to comply with the written description requirement of 35 U.S.C. §112. The '525 patent claimed the benefit of priority of two applications, Application Serial No. 801,343 filed May 27, 1977, and Application Serial No. 805,023 filed June 9, 1977. In the *Fiers* case, party Revel claimed the benefit of priority of an Israeli application filed on November 21, 1979. Thus, the written description inquiry in those case was based on the state of the art at essentially at the "dark ages" of recombinant DNA technology.

The present application has a priority date of July 16, 1998. Much has happened in the development of recombinant DNA technology in the 21 or more years from the time of filing of the applications involved in *Lilly* and *Fiers* and the present application. For example, the technique of polymerase chain reaction (PCR) was invented. Highly efficient cloning and DNA sequencing technology has been developed. Large databases of protein and nucleotide sequences have been compiled. Much of the raw material of the human and other genomes has been sequenced. With these remarkable advances one of skill in the art would recognize that, given the sequence information of SEQ ID NO:1 and SEQ ID NO:3, and the additional extensive detail provided by the subject application, the present inventors were in possession of the claimed polynucleotide variants at the time of filing of this application.

4. Summary

The Office Action failed to base its written description inquiry "on whatever is now claimed." Consequently, the Action did not provide an appropriate analysis of the present claims and how they differ from those found not to satisfy the written description requirement in cases such as *Lilly* and *Fiers*. In particular, the claims of the subject application are fundamentally different from those found invalid in *Lilly* and *Fiers*. The subject matter of the present claims is defined in terms of the chemical structure of SEQ ID NO:1 or SEQ ID NO:3. The courts have stressed that structural features are important factors to consider in a written description analysis of claims to nucleic acids and proteins. In addition, the genus of polynucleotides defined by the present claims is adequately described, as evidenced by Brenner et al and consideration of the claims of the '740 patent involved in *Lilly*. Furthermore, there have been remarkable advances in the state of the art since the *Lilly* and *Fiers* cases, and these advances were given no consideration whatsoever in the position set forth by the Office Action.

Enablement Rejection under 35 U.S.C. § 112, first paragraph

Claims 23 and 26-31 are rejected for allegedly failing to meet the requirements of 35 U.S.C. § 112, first paragraph, on the grounds that the specification does not provide an enabling disclosure commensurate in scope with the claims. In particular, the Examiner alleges that "the specification, while being enabling for DNA molecules encoding the polypeptide of SEQ ID NO:1, does not reasonably provide enablement for DNA molecules encoding a dehydrogenase that is not homologous to SEQ ID NO:11 [sic]. The specification also does not reasonably provide enablement for polynucleotides encoding polynucleotide [sic] having unknown function" (Office Action page 5). Applicants respectfully disagree and traverse the rejections on the following grounds.

The first paragraph of 35 U.S.C. §112 requires that the specification describe how to make and use the claimed subject matter. As set forth in *In re Marzocchi*, 169 USPQ 367, 369 (CCPA 1971):

> The first paragraph of § 112 requires nothing more than objective enablement. How such a teaching is set forth, either by the use of illustrative examples or by broad terminology, is of no importance.

> As a matter of Patent Office practice, then, a specification disclosure which contains a teaching of the manner and process of making and using the invention in terms which correspond in scope to those used in describing and defining the subject matter sought to be patented *must* be take as in compliance with the enabling requirement of the first paragraph of § 112 *unless* there is reason to doubt the objective truth of the statements contained therein which must be relied on for enabling support.

Applicants submit that the disclosure amply enables the claimed invention. Given the sequences of SEQ ID NO:1 and SEQ ID NO:3, one of ordinary skill in the art could readily identify a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence of SEQ ID NO:3, a polynucleotide encoding a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to SEQ ID NO1, or a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to SEQ ID NO:1, using well known methods of sequence analysis without any undue experimentation. In order to expedite prosecution, Applicants have amended claims 21 and 30 to specify that the recited polynucleotide variants encode polypeptides having alcohol dehydrogenase activity. The identification of relevant polynucleotides could be performed by hybridization and/or PCR techniques that were well-known to those skilled in the art at the time the subject application was filed and/or described throughout the Specification of the instant application. See, e.g., page 16, line 18 through page 19, line 16; page 23, lines 7-19; and Example VI at pages 43-44. Thus, one skilled in the art need not make and test vast numbers of polynucleotides. Instead, one skilled in the art need only screen a cDNA library or use appropriate PCR conditions to identify relevant polynucleotides that already exist in nature. The specification also describes expression vectors for the production of the claimed polypeptide variants and fragments, and the construction of fusion proteins (pages 19-24 and Example IX at pages 45-46), and assays to identify polypeptide variants with alcohol dehydrogenase activity (Example X at page 46 and Example XIV at page 48).

Applicants respectfully point out that the claims of the instant application are drawn to **naturally occurring** variants. Thus it is not necessary to screen every conceivable variant which might be made using recombinant methods, as all that is claimed are those variant sequences which are found in nature. Through the process of natural selection, nature will have determined the appropriate sequences.

Further, the Examiner requires working examples (Office Action, page 5). There is no such requirement under the law to provide "working examples." As set forth in *In re Borkowski,* 164 USPQ 642, 645 (CCPA 1970) (footnote omitted):

> However, as we have stated in a number of opinions, a specification need not contain a working example if the invention is otherwise disclosed in such a manner that one skilled in the art will be able to practice it without an undue amount of experimentation.

See also M.P.E.P. 2164.02 as follows:

> Compliance with the enablement requirement of 35 U.S.C. 112, first paragraph, does not turn on whether an example is disclosed. An example may be "working" or "prophetic"... A prophetic example describes an embodiment of the invention based on predicted results rather than work actually conducted or results actually achieved.

Thus, there is no requirement under the law to provide "working examples" of what is claimed. Rather, one looks to whether the specification provides a description of how to make what is claimed. The present specification provides the requisite description.

To enable the claimed invention, Applicants need only disclose information sufficient to permit one of ordinary skill in the art to make and use the invention as claimed, without *undue* experimentation. It is the Examiner's burden to establish that undue experimentation would be necessary to carry out Applicants' invention. *In re Angstadt,* 190 USPQ 214, 219 (CCPA 1976).

Contrary to the standard set forth in *Marzocchi,* the Examiner has failed to provide any ***reasons*** why one would doubt that the guidance provided by the present specification would enable one to make and use the recited polynucleotides. Hence, a *prima facie* case for non-enablement has not been established. For at least the above reasons, withdrawal of the enablement rejection under 35 U.S.C. § 112, first paragraph, is respectfully requested.

Rejections under 35 U.S.C. § 112, second paragraph

Claims 23 and 26-31 are rejected under 35 U.S.C. § 112, second paragraph, as allegedly being indefinite. In particular, it is alleged that "[i]n claim 23, which ultimately depend [sic] from claim 21, the phrases 'biologically active fragment' and 'immunogenic fragment' are unclear because the claim can refer to many polypeptides with different biological and immunogenic activites" (Office Action, page 8). Applicants traverse the rejection on at least the following grounds.

To expedite prosecution, claim 21 c) has been amended to recite that the claimed biologically active fragments have alcohol dehydrogenase activity.

To expedite prosecution, claim 21 d) has been amended to recite that the claimed immunogenic fragments comprise at least 15 contiguous amino acid residues of SEQ ID NO:1 and generate an antibody that specifically binds to a polypeptide comprising an amino acid sequence of SEQ ID NO:1.

These amendments further clarify the intended scope of the claims. For at least the above reasons, withdrawal of the rejections under 35 U.S.C. § 112, second paragraph, is respectfully requested.

Rejections under 35 U.S.C. § 102(b)

Claims 21 and 23 are rejected under 35 U.S.C. § 102(b) as allegedly being anticipated by Bonaldo et al. (1996) Genome Res. 6:791-806 on the grounds that the reference discloses a polynucleotide that comprises at least 60 contiguous nucleotides of SEQ ID NO:3" (Office Action, page 8). Claims 23 and 26-28 are rejected under 35 U.S.C. § 102(b) as allegedly being anticipated by Gabrielli et al. (1995) Eur. J. Biochem. 232:473-477 on the grounds that Gabrielli et al. teach a polynucleotide comprising a fragment of a SCAD protein (Office Action, page 9). Applicants respectfully traverse the rejection.

The publication of Bonaldo et al. Genome Res. 6:791-806 (1996) is irrelevant since it does not disclose the sequences of the instant application, but rather describes methods to "facilitate gene discovery." The sequence CA312519, disclosed by Bonaldo et al., was first made available to the public at NCBI on November 4, 2002 (Please see Exhibit B). The sequences of SEQ ID NO:1 and SEQ ID NO:3 are entitled to the priority date of November 5, 1998. Therefore, the claims are not anticipated by Bonaldo et al., and withdrawal of the rejection under 35 U.S.C. § 102(b) on these grounds is respectfully requested.

Applicants submit that the reference of Gabrielli et al. does not read on claim 21 c), as currently amended, which now recites a biologically active fragment of a polypeptide having an amino acid sequence of SEQ ID NO:1, said fragment having alcohol dehydrogenase activity, nor 21 d), which recites an immunogenic fragment of a polypeptide consisting of an amino acid sequence of SEQ ID NO:1, wherein said fragment comprises at least 15 contiguous amino acid residues of SEQ ID NO:1 and generates an antibody that specifically binds to a polypeptide comprising an amino acid sequence of SEQ ID NO:1. The attention of the Examiner is directed

to Exhibit C, which shows CLUSTALW alignments of the sequences of SEQ ID NO:1 and SEQ ID NO:3 of the instant application with the polypeptide and polynucleotide sequences disclosed by Gabrielli et al. The reference of Gabrielli et al. does not anticipate the claimed polypeptide fragments, nor polynucleotides encoding them, and Applicants respectfully request withdrawal of the rejection under 35 U.S.C. § 102(b).

CONCLUSION

In light of the above amendments and remarks, Applicants submit that the present application is fully in condition for allowance, and request that the Examiner withdraw the outstanding objections/rejections. Early notice to that effect is earnestly solicited.

If the Examiner contemplates other action, or if a telephone conference would expedite allowance of the claims, Applicants invite the Examiner to contact the undersigned at the number listed below.

Applicants believe that no fee is due with this communication. However, if the USPTO determines that a fee is due, the Commissioner is hereby authorized to charge Deposit Account No. **09-0108.**

Respectfully submitted,
INCYTE CORPORATION

Date: March 17, 2004

James M. Verna, Ph.D.
Reg. No. 33,287
Direct Dial Telephone: (650) 845 -5415

Date: March 17, 2004

Jenny Buchbinder
Reg. No. 48,588
Direct Dial Telephone: (650) 843-7212

Customer No.: 27904
3160 Porter Drive
Palo Alto, California 94304
Phone: (650) 855-0555
Fax: (650) 849-8886

Enclosures:
Exhibit A
Exhibit B
Exhibit C

EXHIBIT A
Docket No.: PF-0559 USN
USSN: 09/743,752

SeqServer®
biology in silico

BLAST2 Search Results

Sequences | Help

Retrieval | BLAST2 | FASTA | ClustalW | GCG Assembly | Phrap | Translation
BLAST2 Manual

Confidential -- Property of Incyte Corporation SeqServer Version 4.6 Jan 2002

Program: blastp
Sequence ID(s):
☐ 1240869CD1 vs. genpept137

NCBI-BLASTP 2.0.10 [Aug-26-1999]

Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

Query= 1240869CD1
(278 letters)

Database: genpept137
1,534,369 sequences; 474,463,515 total letters

Searching.....done

Sequences producing significant alignments:		Score (bits)	E Value
☑ g12804321	peroxisomal short-chain alcohol dehydrogenase [Homo	521	e-146
☑ g7023407	unnamed protein product [Homo sapiens]	519	e-146
☑ g11559412	NADPH-dependent retinol dehydrogenase/reductase [Ho	499	e-140
☑ g4105190	peroxisomal short-chain alcohol dehydrogenase [Homo	496	e-139
☑ g17298119	carbonyl reductase/NADP-retinol dehydrogenase [Sus	430	e-119
☑ g11559416	NADPH-dependent retinol dehydrogenase/reductase [Or	430	e-119
☑ g32450664	Unknown (protein for MGC:62554) [Mus musculus]	426	e-118
☑ g19702303	NADPH-dependent retinol dehydrogenase/reductase [Bo	421	e-117
☑ g13097510	Dhrs4 protein [Mus musculus]	417	e-115
☑ g11559414	NADPH-dependent retinol dehydrogenase/reductase [Mu	417	e-115

>g12804321 peroxisomal short-chain alcohol dehydrogenase [Homo sapiens]
Length = 278

Score = 521 bits (1327), Expect = e-146
Identities = 268/278 (96%), Positives = 268/278 (96%)

Query: 1 MHMARLLGLCAWARKSVRMASSRMTRRDPLTNKVALVTASTDGIGFAIARRLAQDRAHVV 60

```
            MH A LLGLCA A  SVRMASS MTRRDPL NKVALVTASTDGIGFAIARRLAQD AHVV
Sbjct: 1    MHKAGLLGLCARAWNSVRMASSGMTRRDPLANKVALVTASTDGIGFAIARRLAQDGAHVV 60

Query: 61   VSSRKQQNVDQAVATLQGEGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVN 120
            VSSRKQQNVDQAVATLQGEGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVN
Sbjct: 61   VSSRKQQNVDQAVATLQGEGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVN 120

Query: 121  PFFGSIMDVTEEVWDKTLDINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFS 180
            PFFGSIMDVTEEVWDKTLDINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFS
Sbjct: 121  PFFGSIMDVTEEVWDKTLDINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFS 180

Query: 181  PYNVSKTALLGLNNTLAIELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIR 240
            PYNVSKTALLGL  TLAIELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIR
Sbjct: 181  PYNVSKTALLGLTKTLAIELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIR 240

Query: 241  RLGEPEDCAGIVSFLCSEDASYITGETVVVGGGTPSRL 278
            RLGEPEDCAGIVSFLCSEDASYITGETVVVGGGTPSRL
Sbjct: 241  RLGEPEDCAGIVSFLCSEDASYITGETVVVGGGTPSRL 278


>g7023407 unnamed protein product [Homo sapiens]
          Length = 278

 Score =  519 bits (1323), Expect = e-146
 Identities = 267/278 (96%), Positives = 267/278 (96%)

Query: 1    MHMARLLGLCAWARKSVRMASSRMTRRDPLTNKVALVTASTDGIGFAIARRLAQDRAHVV 60
            MH A LLGLCA A  SVRMASS MTRRDPL NKVAL TASTDGIGFAIARRLAQD AHVV
Sbjct: 1    MHKAGLLGLCARAWNSVRMASSGMTRRDPLANKVALATASTDGIGFAIARRLAQDGAHVV 60

Query: 61   VSSRKQQNVDQAVATLQGEGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVN 120
            VSSRKQQNVDQAVATLQGEGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVN
Sbjct: 61   VSSRKQQNVDQAVATLQGEGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVN 120

Query: 121  PFFGSIMDVTEEVWDKTLDINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFS 180
            PFFGSIMDVTEEVWDKTLDINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFS
Sbjct: 121  PFFGSIMDVTEEVWDKTLDINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFS 180

Query: 181  PYNVSKTALLGLNNTLAIELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIR 240
            PYNVSKTALLGL  TLAIELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIR
Sbjct: 181  PYNVSKTALLGLTKTLAIELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIR 240

Query: 241  RLGEPEDCAGIVSFLCSEDASYITGETVVVGGGTPSRL 278
            RLGEPEDCAGIVSFLCSEDASYITGETVVVGGGTPSRL
Sbjct: 241  RLGEPEDCAGIVSFLCSEDASYITGETVVVGGGTPSRL 278


>g11559412 NADPH-dependent retinol dehydrogenase/reductase [Homo
           sapiens]
           Length = 260

 Score =  499 bits (1271), Expect = e-140
 Identities = 255/260 (98%), Positives = 255/260 (98%)

Query: 19   MASSRMTRRDPLTNKVALVTASTDGIGFAIARRLAQDRAHVVVSSRKQQNVDQAVATLQG 78
            MASS MTRRDPL NKVALVTASTDGIGFAIARRLAQD AHVVVSSRKQQNVDQAVATLQG
Sbjct: 1    MASSGMTRRDPLANKVALVTASTDGIGFAIARRLAQDGAHVVVSSRKQQNVDQAVATLQG 60

Query: 79   EGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVNPFFGSIMDVTEEVWDKTL 138
            EGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVNPFFGSIMDVTEEVWDKTL
Sbjct: 61   EGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVNPFFGSIMDVTEEVWDKTL 120

Query: 139  DINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFSPYNVSKTALLGLNNTLAI 198
            DINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFSPYNVSKTALLGL  TLAI
Sbjct: 121  DINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFSPYNVSKTALLGLTKTLAI 180

Query: 199  ELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIRRLGEPEDCAGIVSFLCSE 258
```

```
            ELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIRRLGEPEDCAGIVSFLCSE
Sbjct: 181  ELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIRRLGEPEDCAGIVSFLCSE 240

Query: 259  DASYITGETVVVGGGTPSRL 278
            DASYITGETVVVGGGTPSRL
Sbjct: 241  DASYITGETVVVGGGTPSRL 260


>g4105190 peroxisomal short-chain alcohol dehydrogenase [Homo
            sapiens]
            Length = 260

 Score =  496 bits (1263), Expect = e-139
 Identities = 253/260 (97%), Positives = 254/260 (97%)

Query: 19   MASSRMTRRDPLTNKVALVTASTDGIGFAIARRLAQDRAHVVVSSRKQQNVDQAVATLQG 78
            MASS MTRRDPL NKVALVTASTDGIGFAIARRLAQD AHVVVSSRKQQNVDQAVATLQG
Sbjct: 1    MASSGMTRRDPLANKVALVTASTDGIGFAIARRLAQDGAHVVVSSRKQQNVDQAVATLQG 60

Query: 79   EGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVNPFFGSIMDVTEEVWDKTL 138
            EGLSVTGTVCHVGKAEDRERLVA AVKLHGGIDILVSNAAVNPFFGS+MDVTEEVWDKTL
Sbjct: 61   EGLSVTGTVCHVGKAEDRERLVAMAVKLHGGIDILVSNAAVNPFFGSLMDVTEEVWDKTL 120

Query: 139  DINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFSPYNVSKTALLGLNNTLAI 198
            DINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFSPYNVSKTALLGL  TLAI
Sbjct: 121  DINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFSPYNVSKTALLGLTKTLAI 180

Query: 199  ELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIRRLGEPEDCAGIVSFLCSE 258
            ELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIRRLGEPEDCAGIVSFLCSE
Sbjct: 181  ELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIRRLGEPEDCAGIVSFLCSE 240

Query: 259  DASYITGETVVVGGGTPSRL 278
            DASYITGETVVVGGGTPSRL
Sbjct: 241  DASYITGETVVVGGGTPSRL 260


>g17298119 carbonyl reductase/NADP-retinol dehydrogenase [Sus
            scrofa]
            Length = 260

 Score =  430 bits (1093), Expect = e-119
 Identities = 215/260 (82%), Positives = 233/260 (88%)

Query: 19   MASSRMTRRDPLTNKVALVTASTDGIGFAIARRLAQDRAHVVVSSRKQQNVDQAVATLQG 78
            MAS+ + RR PL NKVALVTASTDGIG AIARRLAQD AHVVVSSRKQ+NVD+ VATLQG
Sbjct: 1    MASTGVERRKPLENKVALVTASTDGIGLAIARRLAQDGAHVVVSSRKQENVDRTVATLQG 60

Query: 79   EGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVNPFFGSIMDVTEEVWDKTL 138
            EGLSVTGTVCHVGKAEDRERLVA AV LHGG+DILVSNAAVNPFFG+I+D TEEVWDK L
Sbjct: 61   EGLSVTGTVCHVGKAEDRERLVAMAVNLHGGVDILVSNAAVNPFFGNIIDATEEVWDKIL 120

Query: 139  DINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFSPYNVSKTALLGLNNTLAI 198
             +NVKA  LMTKAVVPEMEKRGGGSV+IVSS+ A+ P P   PYNVSKTALLGL   LA+
Sbjct: 121  HVNVKATVLMTKAVVPEMEKRGGGSVLIVSSVGAYHPFPNLGPYNVSKTALLGLTKNLAV 180

Query: 199  ELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIRRLGEPEDCAGIVSFLCSE 258
            ELAPRNIRVNCLAPGLIKT+FS++LWMDK ++E MKE+LRIRRLG PEDCAGIVSFLCSE
Sbjct: 181  ELAPRNIRVNCLAPGLIKTNFSQVLWMDKARKEYMKESLRIRRLGNPEDCAGIVSFLCSE 240

Query: 259  DASYITGETVVVGGGTPSRL 278
            DASYITGETVVVGGGT SRL
Sbjct: 241  DASYITGETVVVGGGTASRL 260


>g11559416 NADPH-dependent retinol dehydrogenase/reductase
            [Oryctolagus cuniculus]
            Length = 260
```

```
 Score =  430 bits (1093), Expect = e-119
 Identities = 217/260 (83%), Positives = 235/260 (89%)

Query: 19  MASSRMTRRDPLTNKVALVTASTDGIGFAIARRLAQDRAHVVVSSRKQQNVDQAVATLQG 78
           MASS +TRRDPL NKVA+VTASTDGIG AIARRLAQD AHVV+SSRKQQNVD+AVA LQ
Sbjct: 1   MASSGVTRRDPLANKVAIVTASTDGIGLAIARRLAQDGAHVVISSRKQQNVDRAVAALQA 60

Query: 79  EGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVNPFFGSIMDVTEEVWDKTL 138
           EGLSVTGTVCHVGKAEDRERLVATA+ LHGGIDILVSNAAVNPFFG +MDVTEEVWDK L
Sbjct: 61  EGLSVTGTVCHVGKAEDRERLVATALNLHGGIDILVSNAAVNPFFGKLMDVTEEVWDKIL 120

Query: 139 DINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFSPYNVSKTALLGLNNTLAI 198
           DINVKA ALMTKAVVPEMEKRGGGSVVIV+SIAAF+P  G  PYNVSKTAL+GL   LA+
Sbjct: 121 DINVKAMALMTKAVVPEMEKRGGGSVVIVASIAAFNPFSGLGPYNVSKTALVGLTKNLAL 180

Query: 199 ELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIRRLGEPEDCAGIVSFLCSE 258
           ELA +NIRVNCLAPGLIKTSFS+ LW DK +EE++ + LRIRRLG+PE+CAGIVSFLCSE
Sbjct: 181 ELAAQNIRVNCLAPGLIKTSFSKALWEDKAQEENIIQKLRIRRLGKPEECAGIVSFLCSE 240

Query: 259 DASYITGETVVVGGGTPSRL 278
           DASYITGETVVV GG PSRL
Sbjct: 241 DASYITGETVVVAGGAPSRL 260


>g32450664 Unknown (protein for MGC:62554) [Mus musculus]
           Length = 279

 Score =  426 bits (1084), Expect = e-118
 Identities = 216/274 (78%), Positives = 238/274 (86%)

Query: 5   RLLGLCAWARKSVRMASSRMTRRDPLTNKVALVTASTDGIGFAIARRLAQDRAHVVVSSR 64
           RLLG    A  SVRMASS +TRR+PL+NKVALVTASTDGIGFAIARRLA+D AHVVVSSR
Sbjct: 6   RLLGGWTQAWMSVRMASSGLTRRNPLSNKVALVTASTDGIGFAIARRLAEDGAHVVVSSR 65

Query: 65  KQQNVDQAVATLQGEGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVNPFFG 124
           KQQNVD+AVATLQGEGLSVTG VCHVGKAEDRE+L+ TA+K H GIDILVSNAAVNPFFG
Sbjct: 66  KQQNVDRAVATLQGEGLSVTGIVCHVGKAEDREKLITTALKRHRGIDILVSNAAVNPFFG 125

Query: 125 SIMDVTEEVWDKTLDINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFSPYNV 184
           ++MDVTEEVWDK L INV A A+M KAVVPEMEKRGGGSVVIV S+A F+  P   PYNV
Sbjct: 126 NLMDVTEEVWDKVLSINVTATAMMIKAVVPEMEKRGGGSVVIVGSVAGFTRFPSLGPYNV 185

Query: 185 SKTALLGLNNTLAIELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIRRLGE 244
           SKTALLGL    A ELAP+NIRVNCLAPGLIKT FS +LW +K +E+ +KE ++IRRLG+
Sbjct: 186 SKTALLGLTKNFAAELAPKNIRVNCLAPGLIKTRFSSVLWEEKAREDFIKEAMQIRRLGK 245

Query: 245 PEDCAGIVSFLCSEDASYITGETVVVGGGTPSRL 278
           PEDCAGIVSFLCSEDASYI GETVVVGGGTPSRL
Sbjct: 246 PEDCAGIVSFLCSEDASYINGETVVVGGGTPSRL 279


>g19702303 NADPH-dependent retinol dehydrogenase/reductase [Bos
           taurus]
           Length = 260

 Score =  421 bits (1072), Expect = e-117
 Identities = 210/260 (80%), Positives = 234/260 (89%)

Query: 19  MASSRMTRRDPLTNKVALVTASTDGIGFAIARRLAQDRAHVVVSSRKQQNVDQAVATLQG 78
           MAS  M RR+PL NKVALVTASTDGIGFAIARRLAQD AHVVVSSRKQQNVD+AVATL+G
Sbjct: 1   MASCGMARRNPLDNKVALVTASTDGIGFAIARRLAQDGAHVVVSSRKQQNVDRAVATLKG 60

Query: 79  EGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVNPFFGSIMDVTEEVWDKTL 138
           EGLSVTGTVCHVGKAEDRERLVATAVKLHGG+DIL+SNAAV+PFFGS+MDV EEVWDK L
Sbjct: 61  EGLSVTGTVCHVGKAEDRERLVATAVKLHGGVDILISNAAVSPFFGSLMDVPEEVWDKIL 120
```

```
Query: 139 DINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFSPYNVSKTALLGLNNTLAI 198
           D+NVKA AL+TKAVVPEM KRGGGS+VIVSSIAA+SP P   PYNVSKTALLGL   LA+
Sbjct: 121 DVNVKATALLTKAVVPEMAKRGGGSIVIVSSIAAYSPFPSLGPYNVSKTALLGLTKNLAL 180

Query: 199 ELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIRRLGEPEDCAGIVSFLCSE 258
           ELA  N+RVNCLAPGLI+TSFSR+LW D  ++ES+K T +I+R+G+PE+CAGIVSFLCSE
Sbjct: 181 ELAESNVRVNCLAPGLIRTSFSRVLWEDPARQESIKATFQIKRIGKPEECAGIVSFLCSE 240

Query: 259 DASYITGETVVVGGGTPSRL 278
           DASYITGETVVV GG+ S L
Sbjct: 241 DASYITGETVVVAGGSLSHL 260


>g13097510 Dhrs4 protein [Mus musculus]
           Length = 260

 Score =  417 bits (1060), Expect = e-115
 Identities = 208/260 (80%), Positives = 230/260 (88%)

Query: 19  MASSRMTRRDPLTNKVALVTASTDGIGFAIARRLAQDRAHVVVSSRKQQNVDQAVATLQG 78
           MASS +TRR+PL+NKVALVTASTDGIGFAIARRLA+D AHVVVSSRKQQNVD+AVATLQG
Sbjct: 1   MASSGLTRRNPLSNKVALVTASTDGIGFAIARRLAEDGAHVVVSSRKQQNVDRAVATLQG 60

Query: 79  EGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVNPFFGSIMDVTEEVWDKTL 138
           EGLSVTG VCHVGKAEDRE+L+ TA+K H GIDILVSNAAVNPFFG++MDVTEEVWDK L
Sbjct: 61  EGLSVTGIVCHVGKAEDREKLITTALKRHRGIDILVSNAAVNPFFGNLMDVTEEVWDKVL 120

Query: 139 DINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFSPYNVSKTALLGLNNTLAI 198
            INV A A+M KAVVPEMEKRGGGSVVIV S+A F+  P   PYNVSKTALLGL    A
Sbjct: 121 SINVTATAMMIKAVVPEMEKRGGGSVVIVGSVAGFTRFPSLGPYNVSKTALLGLTKNFAA 180

Query: 199 ELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIRRLGEPEDCAGIVSFLCSE 258
           ELAP+NIRVNCLAPGLIKT FS +LW +K +E+ +KE ++IRRLG+PEDCAGIVSFLCSE
Sbjct: 181 ELAPKNIRVNCLAPGLIKTRFSSVLWEEKAREDFIKEAMQIRRLGKPEDCAGIVSFLCSE 240

Query: 259 DASYITGETVVVGGGTPSRL 278
           DASYI GETVVVGGGTPSRL
Sbjct: 241 DASYINGETVVVGGGTPSRL 260


>g11559414 NADPH-dependent retinol dehydrogenase/reductase [Mus
           musculus]
           Length = 260

 Score =  417 bits (1060), Expect = e-115
 Identities = 208/260 (80%), Positives = 230/260 (88%)

Query: 19  MASSRMTRRDPLTNKVALVTASTDGIGFAIARRLAQDRAHVVVSSRKQQNVDQAVATLQG 78
           MASS +TRR+PL+NKVALVTASTDGIGFAIARRLA+D AHVVVSSRKQQNVD+AVATLQG
Sbjct: 1   MASSGLTRRNPLSNKVALVTASTDGIGFAIARRLAEDGAHVVVSSRKQQNVDRAVATLQG 60

Query: 79  EGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVNPFFGSIMDVTEEVWDKTL 138
           EGLSVTG VCHVGKAEDRE+L+ TA+K H GIDILVSNAAVNPFFG++MDVTEEVWDK L
Sbjct: 61  EGLSVTGIVCHVGKAEDREKLITTALKRHQGIDILVSNAAVNPFFGNLMDVTEEVWDKVL 120

Query: 139 DINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFSPYNVSKTALLGLNNTLAI 198
            INV A A+M KAVVPEMEKRGGGSVVIV S+A F+  P   PYNVSKTALLGL    A
Sbjct: 121 SINVTATAMMIKAVVPEMEKRGGGSVVIVGSVAGFTRFPSLGPYNVSKTALLGLTKNFAA 180

Query: 199 ELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIRRLGEPEDCAGIVSFLCSE 258
           ELAP+NIRVNCLAPGLIKT FS +LW +K +E+ +KE ++IRRLG+PEDCAGIVSFLCSE
Sbjct: 181 ELAPKNIRVNCLAPGLIKTRFSSVLWEEKAREDFIKEAMQIRRLGKPEDCAGIVSFLCSE 240

Query: 259 DASYITGETVVVGGGTPSRL 278
           DASYI GETVVVGGGTPSRL
Sbjct: 241 DASYINGETVVVGGGTPSRL 260
```

```
  Database: genpept137
    Posted date:  Sep 11, 2003 11:22 AM
  Number of letters in database: 474,463,515
  Number of sequences in database:  1,534,369

Lambda     K      H
   0.318    0.133    0.378

Gapped
Lambda     K      H
   0.270   0.0470    0.230


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 242961409
Number of Sequences: 1534369
Number of extensions: 9442916
Number of successful extensions: 37127
Number of sequences better than 10.0: 4670
Number of HSP's better than 10.0 without gapping: 3208
Number of HSP's successfully gapped in prelim test: 1462
Number of HSP's that attempted gapping in prelim test: 26620
Number of HSP's gapped (non-prelim): 4858
length of query: 278
length of database: 474,463,515
effective HSP length: 61
effective length of query: 217
effective length of database: 380,867,006
effective search space: 82648140302
effective search space used: 82648140302
T: 11
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.8 bits)
X3: 64 (24.9 bits)
S1: 41 (21.7 bits)
```

Submit sequences to: BLAST2 Submit Reset

IncyteGenomics

EXHIBIT B
Docket No.: PF-0559 USN
USSN: 09/743,752

NCBI

Sequence Revision History

Revision history for 24530617

GI	Version	Update Date	Status
24530617	1	Nov 4 2002 5:05	Live

Accession CA312519 was first seen at NCBI on Nov 4 2002 5:05

Entrez

Search for Genes
LocusLink provides curated information for human, fruit fly, mouse, rat, and zebrafish

Help|FAQ

Batch Entrez: Upload a file of GI or accession numbers to retrieve protein or nucleotide sequences

Check sequence revision history

How to create WWW links to Entrez

LinkOut

Cubby

Related resources

BLAST

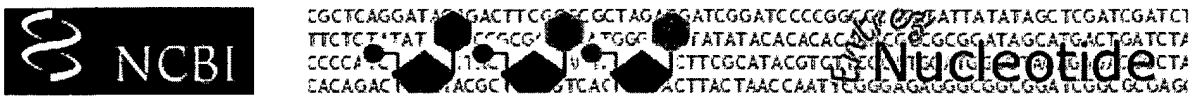
Reference sequence project

LocusLink

Clusters of orthologous groups

Protein reviews on the web

Disclaimer | Write to the Help Desk
NCBI | NLM|NIH



1: CA312519. UI-CF-FN0-afk-l-1...[gi:24530617]

IDENTIFIERS

dbEST Id: **14963301**
EST name: UI-CF-FN0-afk-l-13-0-UI.s1
GenBank Acc: CA312519
GenBank gi: 24530617

CLONE INFO
Clone Id: UI-CF-FN0-afk-l-13-0-UI (3')
Source: University of Iowa and Cystic Fibrosis Foundation - Cystic Fibrosis Project
Id as DNA: UI-CF-FN0-afk-l-13-0-UI.s1
Id in host: UI-CF-FN0-afk-l-13-0-UI
DNA type: cDNA

PRIMERS
Sequencing: M13 FORWARD
PolyA Tail: yes

SEQUENCE

TTTTTTTTTTTTTTTTTTTCAAAGCCGCGCAATCATCTGCATTTTATTTGCACCTCATTTG
CAAGTTGTTAATTTGCAACTCTGCTCCTTCCACTCCAGGTTCCTTCTCTCCCATCACCCC
TCAGATTCCCCAGTGGCCCAGGAAAAAATATCTTGGTCTTTGCCAAGGTAGACTCAGCCT
TGTCAGCAGGCCTGTCCTGTGTTCTCAGGGGAGGCCTTTACCCAAGGCCACAACAACAGC
AGGAATCCCGAGTAAGACGCCACCTTGACGGCAGGGAAGGCTGGATCTTTTCACAGGGCA
GAACTGATTTGATGAGGTGAACAGTAAGGTGAGCAGAGGTGGGAAAGGCCAGTGGGTGAA
TGCAGGAACAGCACCAGGAGCTAGAGCCCAACTCTGGCCTGTGGGCTGTCTCCCGGTCCT
CAGAGGCGGGACGGGGTTCCTCCACCCACCACCACTGTTTCCCCAGTGATGTAGCTGGCA
TCTTCAGAGCACAGGAAAGACACGATGCCAGCACAATCCTCTGGCTCGCCTAACCTTCTT
ATCCGCAGGGTTTCTTTCATGCTTTCCTCTTTTTCCTTGTCCATCCAGAGCATCCTGCTG
AAGCTAGTCTTGATAAGTCCAGGTGCTAGGCAGTTCACCCTAATGTTCCTTGGGGCCAGC
TCTATGGCCAGGGTCTTGGTCAGGCCCAGCAAGGCTGTTTTACTGACATTGNTAGGACTG
AAGCCAGGAGATGGACTG

Entry Created: Nov 4 2002
Last Updated: Nov 4 2002

COMMENTS

Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) or from Open Biosystems (www.openbiosystems.com).

LIBRARY
Lib Name: UI-CF-FN0
Organism: Homo sapiens
Tag Seq: CTGCTCAGGT
Tag Tissue: Human Lung Epithelial Cell Lines untreated LPS 6hr to LPS 24h
Tag Lib: UI-CF-FN0
Organ: Lung

Tissue type:	Human Lung Epithelial cells
Lab host:	DH10B (Life Technologies) (T1 phage resistant)
Vector:	pT7T3-Pac (Pharmacia) with a modified polylinker
R. Site 1:	EcoR I
R. Site 2:	Not I
Description:	UI-CF-FN0 is a subtracted cDNA library derived from two normalized Human lung epithelial cell libraries (EN1 and DU1) The library was subtracted according to according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bento-soares@uiowa.edu

SUBMITTER

Name:	McCray, PB
Lab:	McCray Lab
Institution:	University of Iowa
Address:	2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel:	319 356 4866
Fax:	319 356 7171
E-mail:	paul-mccray@uiowa.edu

CITATIONS

Medline UID:	97044477
Title:	Normalization and subtraction: two approaches to facilitate gene discovery
Authors:	Bonaldo,M.F., Lennon,G., Soares,M.B.
Citation:	Genome Res. 6 (9): 791-806 1996

MAP DATA

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

Jan 29 2004 15:38:25

SeqServer®
biology in silico

ClustalW Results

EXHIBIT C
Docket No.: PF-0559 USN
USSN: 09/743,752

Sequences | Help

Retrieval | BLAST2 | FASTA | ClustalW | GCG Assembly | Phrap | Translation

Confidential -- Property of Incyte Corporation SeqServer Version 4.6 Jan 2002

- g1079566
- 1240869CD1

CLUSTAL W (1.7) Multiple Sequence Alignments

```
Sequence format is Pearson
Sequence 1: g1079566        280 aa
Sequence 2: 1240869CD1      278 aa
Start of Pairwise alignments
Aligning...
Sequences (1:2) Aligned. Score:  56
Start of Multiple Alignment
There are 1 groups
Aligning...
Group 1: Sequences:   2      Score:2822
Alignment Score 975
CLUSTAL-Alignment file created  [baa7VaW8k.aln]
CLUSTAL W (1.7) multiple sequence alignment


g1079566        MLSAVARGYQGWFHPCARLSVRMSSTGIDRKGVLANRVAVVTGSTSGIGFAIARRLARDG
1240869CD1      MHMARLLGLCAW----ARKSVRMASSRMTRRDPLTNKVALVTASTDGIGFAIARRLAQDR
                *  *   *  .*    ** ****:*: : *:. *:*:**:**.**.***********:*

g1079566        AHVVISSRKQQNVDRAMAKLQGEGLSVAGIVCHVGKAEDREQLVAKALEHCGGVDFLVCS
1240869CD1      AHVVVSSRKQQNVDQAVATLQGEGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSN
                ****:*********:*:*.********:* ***********:***.*::  **:*:**..

g1079566        AGVNPLVGSTLGTSEQIWDKILSVNVKSPALLLSQLLPYMENRRG-AVILVSSIAAYNPV
1240869CD1      AAVNPFFGSIMDVTEEVWDKTLDINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPS
                *.***:.** :..:*::*** *.:***:***: . ::* **:* * :*::******:.*

g1079566        VALGVYNVSKTALLGLTRTLALELAPKDIRVNCVVPGIIKTDFSKVFHGNESLWKNFKEH
1240869CD1      PGFSPYNVSKTALLGLNNTLAIELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKET
                 .:. ***********..***:****::*****:.**:***.**:::  ::.  :.:**

g1079566        HQLQRIGESEDCAGIVSFLCSPDASYVNGENIAVAG-YSTRL
1240869CD1      LRIRRLGEPEDCAGIVSFLCSEDASYITGETVVVGGGTPSRL
                 :::*:**.************ ****:.**.:.*.*  .:**
```

SeqServer®
biology in silico

ClustalW Results

Sequences | Help

Retrieval | BLAST2 | FASTA | ClustalW | GCG Assembly | Phrap | Translation

Confidential -- Property of Incyte Corporation SeqServer Version 4.6 Jan 2002

- g1079565
- 1240869CB1

CLUSTAL W (1.7) Multiple Sequence Alignments

```
Sequence format is Pearson
Sequence 1: g1079565       1442 bp
Sequence 2: 1240869CB1     1280 bp
Start of Pairwise alignments
Aligning...
Sequences (1:2) Aligned. Score:  44
Start of Multiple Alignment
There are 1 groups
Aligning...
Group 1: Sequences:   2      Score:12085
Alignment Score 4402
CLUSTAL-Alignment file created  [baa81aqU1.aln]
CLUSTAL W (1.7) multiple sequence alignment


g1079565        GGTTCCCTTCCACGCTGTGAAGCTTTGTTCTTTTGGTCTTCATGATAAATCTTGCTGCTG
1240869CB1     ------------------------------------------------------------


g1079565        CTCACTCGTTGGGTCCGTGCCACCTTTAAGAGCTGTAACACTCACCGCGAAGGTCTGCAA
1240869CB1     ------------------------------------------------------------


g1079565        CTTCACTCCTGGGGCCAGCAAGACCACGAATGCACCGAGAGGAATGAACAACTCTGGACA
1240869CB1     ------------------------------------------------------------


g1079565        CACCATCTTTAAGAACCGTAATACTCACCGCAAGGGTCTGCAACTTCATTCTTGAAGTCA
1240869CB1     ------------------------------------------------------------


g1079565        GTGAGGCCAAGAACCCATCAATTCCGTACACATTTTGGTGACTTTGAAGAGACTGTCACC
1240869CB1     ------------------------------------------------------------


g1079565        TATCACCAAGTGGTGAGACTATTGCCAAGCAGTGAGACTATTGCCAAGTGGTGAGACCAT
1240869CB1     ------------------------------------------------------------
```

```
g1079565        CACCAAGCGGTGAGACTATCACCTATCGCCAAGTGGCCTGATTCAGCAGGAAGCATCTCA
1240869CB1     ----------------------------------------------GAGTGGAACCCA
                                                             *   * * * **

g1079565        GACACCAACCACTATGCTGTCAGCAGTTGCCCGGGGCTACCAGGGCTGGTTTCATCCCTG
1240869CB1     GACTT--GCTGGTCTGATCCATGCACATGGCCAGG-CTGCTAGGCCTCTGTGC----CTG
                ***     *   * ** *    ***  ** ** ** ** * *** **   * *    ***

g1079565        TGCTAGGCTTTCTGTGAGGATGAGCAGCACCGGGATAGACAGGAAGGGCGTCCTGGCTAA
1240869CB1     GGCACGGAAGTCGGTGCGGATGGCCAGCTCCAGGATGACCCGCCGGGACCCGCTCACAAA
                **  **   ** *** *****  **** ** ****   * *   ** *   **  * **

g1079565        CCGGGTAGCCGTGGTCACGGGGTCCACCAGTGGGATCGGCTTTGCCATCGCCCGACGTCT
1240869CB1     TAAGGTGGCCCTGGTAACGGCCTCCACCGACGGGATCGGCTTCGCCATCGCCCGGCGTTT
                   *** *** **** ****  ******   ********** ********** *** *

g1079565        GGCCCGGGACGGGGCCCACGTGGTCATCAGCAGCCGGAAGCAGCAGAACGTGGACCGGGC
1240869CB1     GGCCCAGGACAGGGCCCACGTGGTCGTCAGCAGCCGGAAGCAGCAGAATGTGGACCAGGC
                ***** **** ************** ********************** ******* ***

g1079565        CATGGCCAAGCTGCAGGGGGAGGGGCTGAGTGTGGCGGGCATTGTGTGCCACGTGGGGAA
1240869CB1     GGTGGCCACGCTGCAGGGGGAGGGGCTGAGCGTGACGGGCACCGTGTGCCATGTGGGGAA
                  ****** ********************** *** ******  ******** ********

g1079565        GGCTGAGGACCGGGAGCAGCTGGTGGCCAAGGCCCTGGAGCACTGTGGGGGCGTCGACTT
1240869CB1     GGCGGAGGACCGGGAGCGGCTGGTGGCCACGGCTGTGAAGCTTCATGGAGGTATCGATAT
                *** ************* *********** ***  ** ***    *** **  **** *

g1079565        CCTGGTGTGCAGCGCAGGGGTCAACCCTCTGGTAGGGAGCACTCTGGGGACCAGTGAGCA
1240869CB1     CCTAGTCTCCAATGCTGCTGTCAACCCTTTCTTTGGAAGCATAATGGATGTCACTGAGGA
                *** ** * **  ** *  ********* *  * ** ****   ***    ** **** *

g1079565        GATCTGGGACAAGATCCTAAGTGTGAACGTGAAGTCCCCAGCCCTGCTG-CTGAGCCAGT
1240869CB1     GGTGTGGGACAAGACTCTGGACATTAATGTGAAGGCCCCAGCCCTGATGACAAAGGCAGT
                * * **********  **     * ** ****** ********** ** *  ** ****

g1079565        TGCTGCCCTACATGGAGAAC--AGGAGG-GGTGCTGTCATCCTGGTCTCTTCCATTGCAG
1240869CB1     GG-TGCCAGAAATGGAGAAACGAGGAGGCGGCTCAGTGGTGATCGTGTCTTCCATAGCAG
                 * ****  * ********   ****** **  * **  *  * ** ******** ****

g1079565        CTTATAATCCAG-TAGTGGCGCTGGGTGTCTACAATGTCAGCAAGACAGCGCTGCTGGGT
1240869CB1     CCTTCAGTCCATCTCCTGGCT-TCAGTCCTTACAATGTCAGTAAAACAGCCTTGCTGGGC
                * *  * ****  *  ****  *  **   ********** ** *****  *******

g1079565        CTCACTAGAACACTGGCATTGGAGCTGGCCCCCAAGGACATCCGGGTAAACTGCGTGGTT
1240869CB1     CTCAACAATACCCTGGCCATAGAGCTGGCCCCAAGGAACATTAGGGTGAACTGCCTAGCA
                ****  *  ** *****  * ********** * * ****  **** ****** * *

g1079565        CCAGGAATTATAAAAACTGACTTCAGCAAAGTGTTTCATGGGAATGAGTCTCTCTGGAAG
1240869CB1     CCTGGACTTATCAAGACTAGCTTCAGCAGGATGCTCTGGATGGACAAGGAAAAAGAGGAA
                ** *** **** ** ***  ********   ** *      * *  **        * *

g1079565        AACTTCAAGGAACATCATCAGCTGCAGAGGATTGGGGAGTCAGAGGACTGTGCAGGAATC
1240869CB1     AGCATGAAAGAAACCCTGCGGATAAGAAGGTTAGGCGAGCCAGAGGATTGTGCTGGCATC
                * * * ** ***   *  * * *    *** * ** *** ******* ***** ** ***

g1079565        GTGTCCTTCCTGTGCTCTCCAGATGCCAGCTACGTCAACGGGGAGAACATTGCGGTGGC-
1240869CB1     GTGTCTTTCCTGTGCTCTGAAGATGCCAGCTACATCACTGGGGAAACAGTGGTGGTGGGT
                ***** ***********  ************* ***  ***** *   * * *****

g1079565        --AGGCTACTCCACTCGGCTCTGAGAGGAGTGGG---GGCGGCTGCGTAGCTGTGGTCCC
1240869CB1     GGAGGAACCCCGTCCCGCCTCTGAGGACCGGGAGACAGCCCACAGGCCAGAGTTGGGCTC
                  ***   * *  * ** *******    * * *   * *  * *   **   *** * *

g1079565        -AGC-CCAGGAGCC--TGAGGGGGTGTCTAGGTGATCATTTGGATCTGGAGCAGAGTCTG
```

```
1240869CB1      TAGCTCCTGGTGCTGTTCCTGCATTCACCCACTGGCCTTTCCCACCTCT-GCTCACCTTA
                 *** ** ** **   *   *   *  *    **  * **   * **   **  *   *

g1079565        CCATTCTGCCAGACTAGCAATTTGGGGGCTTACTCATGCT--AGGCTT----GAGG--AA
1240869CB1      CTGTTCACCTCATCAAATCAGTTCTGCCCTGTGAAAAGATCCAGCCTTCCCTGCCGTCAA
                *  ***  *    * *   * **  *  **     * * *  ** ***    *  *  **

g1079565        GAAGAAAAACGCTTCGGCATTC-TCC----------------------------------
1240869CB1      GGTGGCGTCTTACTCGGGATTCCTGCTGTTGTTGTGGCCTTGGGTAAAGGCCTCCCCTGA
                *  *         **** **** * *

g1079565        ------------------------------------------------------------
1240869CB1      GAACACAGGACAGGCCTGCTGACAAGGCTGAGTCTACCTTGGCAAAGACCAAGATATTTT

g1079565        ------------------------------------------------------------
1240869CB1      TTCCTGGGCCACTGGGGAATCTGAGGGGTGATGGGAGAGAAGGAACCTGGAGTGGAAGGA

g1079565        ------------------------------------------------------------
1240869CB1      GCAGAGTTGCAAATTAACAACTTGCAAATGAGGTGCAAATAAAATGCAGATGATTGCGCG

g1079565        ------------------
1240869CB1      GCTTTGAAAAAAAAAAAA
```

Submit sequences to: BLAST2 Submit Reset

IncyteGenomics


```
LOCUS       HSU31875                1442 bp    mRNA    linear   PRI 29-NOV-1995
DEFINITION  Human Hep27 protein mRNA, complete cds.
ACCESSION   U31875
VERSION     U31875.1  GI:1079565
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1442)
  AUTHORS   Gabrielli,F., Donadel,G., Bensi,G., Heguy,A. and Melli,M.
  TITLE     A nuclear protein, synthesized in growth-arrested human
            hepatoblastoma cells, is a novel member of the short-chain alcohol
            dehydrogenase family
  JOURNAL   Eur. J. Biochem. 232 (2), 473-477 (1995)
  MEDLINE   96035881
   PUBMED   7556196
REFERENCE   2  (bases 1 to 1442)
  AUTHORS   Gabrielli,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-1995) Franco Gabrielli, Physiology and
            Biochemistry, University of Pisa, Via Roma 55, Pisa 56126, Italy
FEATURES             Location/Qualifiers
     source          1..1442
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /cell_line="hepetoblastoma HepG2"
     CDS             434..1276
                     /note="similar to Streptomyces violaceoruber granaticin
                     polyketide synthase putative ketoacyl, Swiss-Prot
                     Accession Number P16542"
                     /codon_start=1
                     /product="Hep27 protein"
                     /protein_id="AAA82048.1"
                     /db_xref="GI:1079566"
                     /translation="MLSAVARGYQGWFHPCARLSVRMSSTGIDRKGVLANRVAVVTGS
                     TSGIGFAIARRLARDGAHVVISSRKQQNVDRAMAKLQGEGLSVAGIVCHVGKAEDREQ
                     LVAKALEHCGGVDFLVCSAGVNPLVGSTLGTSEQIWDKILSVNVKSPALLLSQLLPYM
                     ENRRGAVILVSSIAAYNPVVALGVYNVSKTALLGLTRTLALELAPKDIRVNCVVPGII
                     KTDFSKVFHGNESLWKNFKEHHQLQRIGESEDCAGIVSFLCSPDASYVNGENIAVAGY
                     STRL"
ORIGIN
        1 ggttcccttc cacgctgtga agctttgttc ttttggtctt catgataaat cttgctgctg
       61 ctcactcgtt gggtccgtgc cacctttaag agctgtaaca ctcaccgcga aggtctgcaa
      121 cttcactcct ggggccagca agaccacgaa tgcaccgaga ggaatgaaca actctggaca
      181 caccatcttt aagaaccgta atactcaccg caagggtctg caacttcatt cttgaagtca
      241 gtgaggccaa gaacccatca attccgtaca cattttggtg actttgaaga gactgtcacc
      301 tatcaccaag tggtgagact attgccaagc agtgagacta ttgccaagtg gtgagaccat
      361 caccaagcgg tgagactatc acctatcgcc aagtggcctg attcagcagg aagcatctca
      421 gacaccaacc actatgctgt cagcagttgc ccggggctac cagggctggt ttcatccctg
      481 tgctaggctt tctgtgagga tgagcagcac cgggatagac aggaagggcg tcctggctaa
      541 ccgggtagcc gtggtcacgg ggtccaccag tgggatcggc tttgccatcg cccgacgtct
      601 ggcccgggac ggggcccacg tggtcatcag cagccggaag cagcagaacg tggaccgggc
      661 catggccaag ctgcaggggg aggggctgag tgtggcgggc attgtgtgcc acgtggggaa
      721 ggctgaggac cgggagcagc tggtggccaa ggccctggag cactgtgggg gcgtcgactt
      781 cctggtgtgc agcgcagggg tcaaccctct ggtagggagc actctgggga ccagtgagca
      841 gatctgggac aagatcctaa gtgtgaacgt gaagtcccca gccctgctgc tgagccagtt
      901 gctgccctac atggagaaca ggaggggtgc tgtcatcctg gtctcttcca ttgcagctta
      961 taatccagta gtggcgctgg gtgtctacaa tgtcagcaag acagcgctgc tgggtctcac
     1021 tagaacactg gcattggagc tggcccccaa ggacatccgg gtaaactgcg tggttccagg
     1081 aattataaaa actgacttca gcaaagtgtt tcatgggaat gagtctctct ggaagaactt
```

```
     1141 caaggaacat catcagctgc agaggattgg ggagtcagag gactgtgcag gaatcgtgtc
     1201 cttcctgtgc tctccagatg ccagctacgt caacggggag aacattgcgg tggcaggcta
     1261 ctccactcgg ctctgagagg agtgggggcg gctgcgtagc tgtggtccca gcccaggagc
     1321 ctgagggggt gtctaggtga tcatttggat ctggagcaga gtctgccatt ctgccagact
     1381 agcaatttgg gggcttactc atgctaggct tgaggaagaa gaaaaacgct tcggcattct
     1441 cc
//
```

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

Jan 29 2004 15:38:25